How to make the ultimate goal of energy-efficient data centers a reality

Julita Corbalan
julita.corbalan@eas4dc.com
julita.corbalan@bsc.es

EAR
Energy Aware Runtime

Why do we need energy efficient Data Centers

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

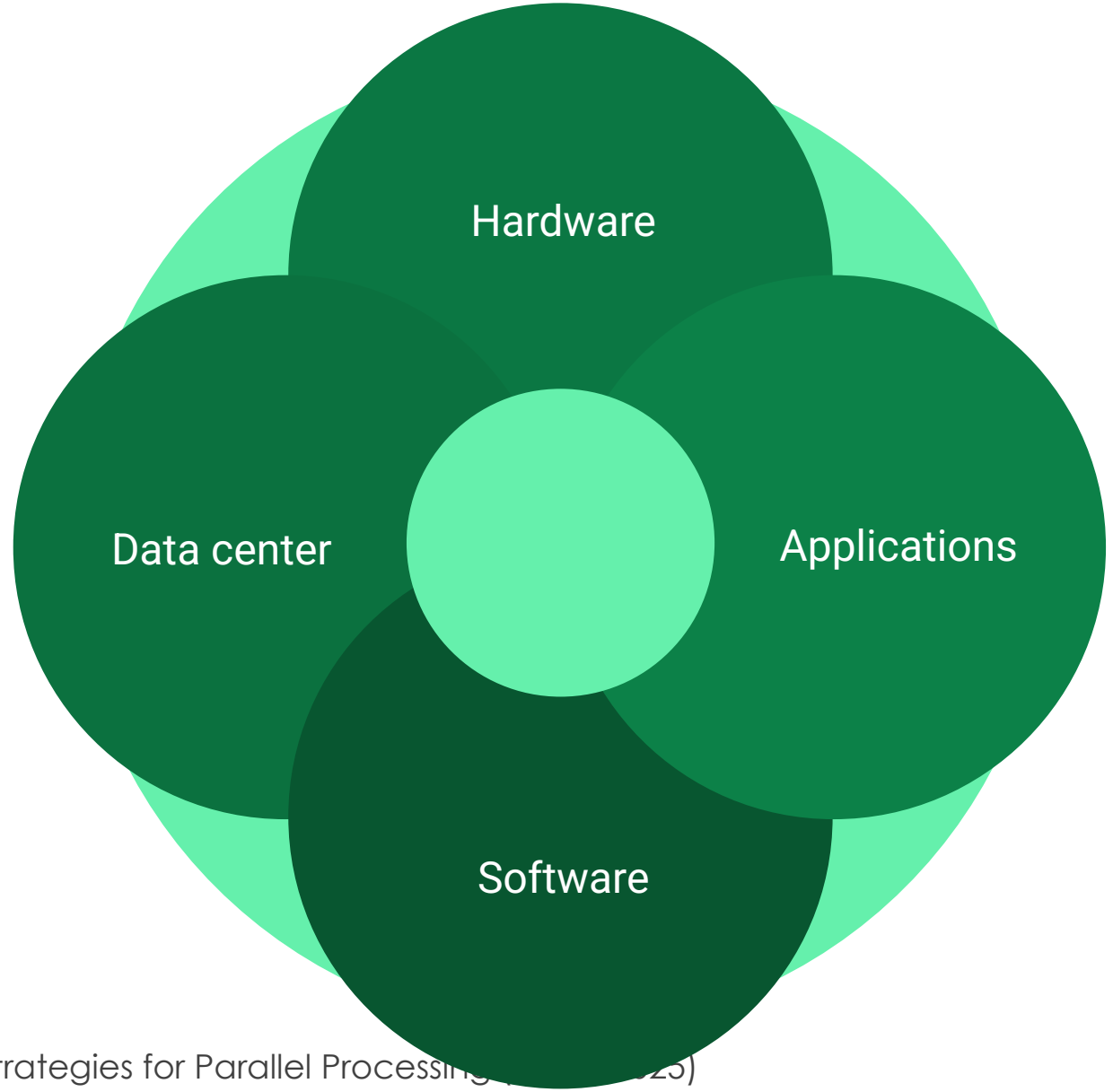
Power is a limited resource

- Massive increase in power requirements
  Real World Example:
  - Initial power available from public grid 8 MW
  - 100k H100 GPUs @700W each 70 MW
- Total data center power estimate ~150 MW
- Based on some sources…
  - 2%- 3% of annual increase if we have a “eco” mind improving hardware and software technologies
  - 5% - 7% of annual increase if the utilization of AI , edge etc increases its utilization without the equivalent improvement in the technologies efficiency

Is it only a matter of money? Do the market will regulate the “efficiency”?

- No (in my opinion)
- The market will regulate what is economically efficiency BUT we cannot destroy the planet for money
- Can we avoid people using data centers , AI chats, cryptocurrencies, etc? Can we control?
  - Buf, very difficult
- What can we do then???
  - **DO OUR BEST !!!**

Energy “efficiency” in Data centers



Data center efficiency

- Metric
  - Energy = Time x Power (Joules or KW/H)
  - Co2
    - CO2 = Energy (KWH) x Carbon Intensity x PUE
    - Carbon intensity factor depends on the country , source of energy etc
    - PUE depends on the DC
- **Data center optimizations**
  - Better cooling infrastructure
  - Use green energies
  - Implement heat reuse
  - etc

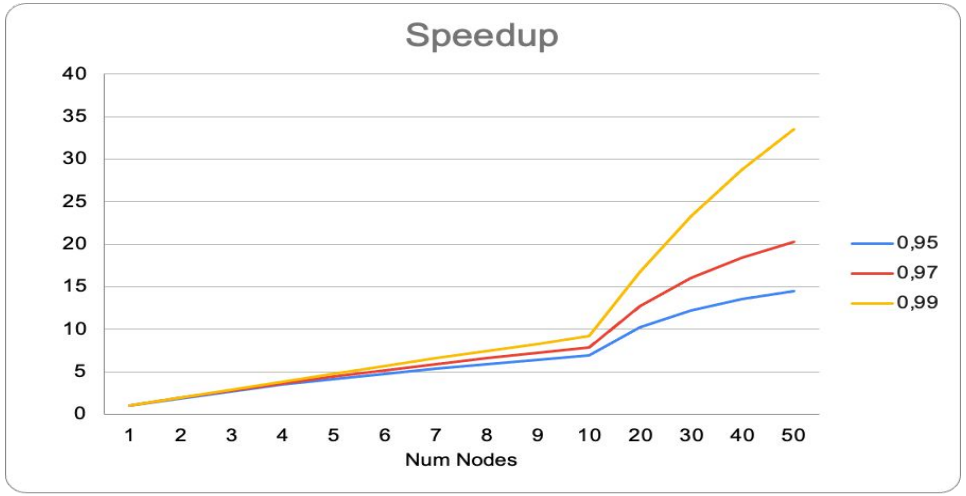
Hardware efficiency

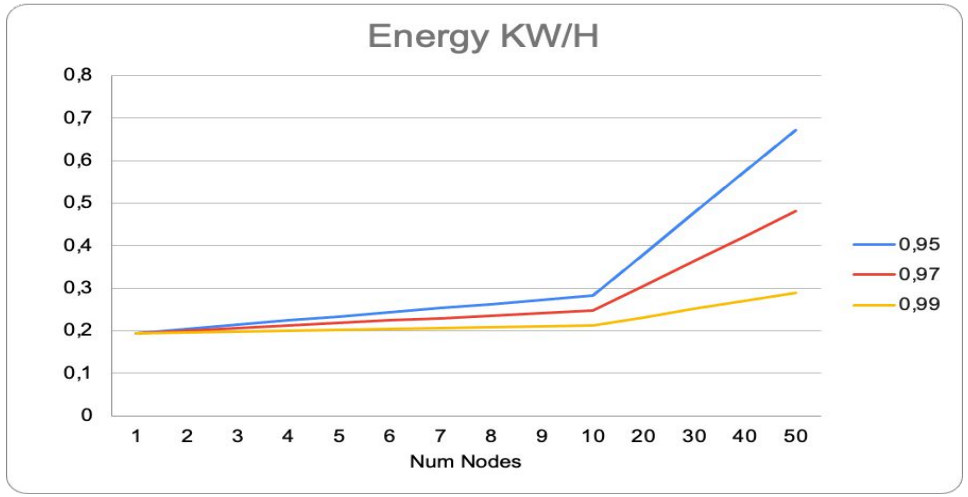
- For many years they have been focusing on increasing performance
- **Performance has been measured in FLOPS, using CPU intensive applications (HPL)**
- Last years, **energy efficiency** has been more and more relevant
  - Metric GFLOPS/Watt
- However, event they are more energy efficient, they are still very high power consuming

- They have "**High Performance Computing**" mind-set (I like the name, it's not mine)
  - We build the fastest machines
  - We break application performance records
  - No Limits

Application efficiency

- Including application developers
- They mostly care about "**Time to solution**"
- They have a very important role in energy optimization
  - Using optimized libraries
  - **Doing a responsible utilization of resources**





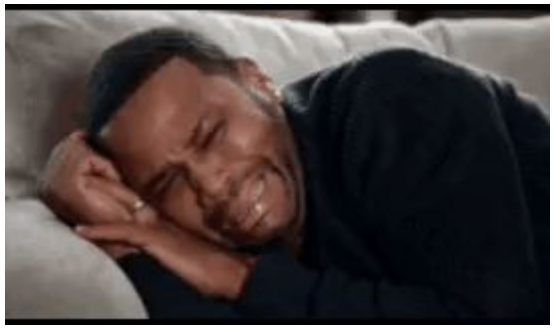
Application efficiency: Performance

- There are lot of CPU/GPU profiling tools
  - Some of them more intrusive, affecting the analysis
  - Some of them very CPU/GPU vendor (or hardware specific)
  - Some at the hardware level → Not valid for regular users
  - Others like perf, papi etc are valid at the process level

Application efficiency: Power/Energy

- What about power
  - Much less expertise and tools since power is mostly a privileged metric
  - Many different APIs, more standard at the CPU level but anyway complex
  - **Different sources of power**
- What about new use cases
  - Workflows
  - Node sharing
  - Virtualization

And we have not commented yet about evaluation metrics!



System software for energy efficient Data centers

- System software tools and service helping application developers to improve their codes
  - Static optimization
    - Energy
    - **Resources (num nodes)… and then energy**
  - Dynamic optimization
    - Energy
    - **Resource (power/frequency) ….. and then energy**
- From "**High Performance Computing**" to "**High Efficient Computing**" mind-set
  - We build the most efficient machines
  - We break application efficiency records
  - We live in a world of limited resources

Stakeholders perspective on Energy Optimization

- Users
  - Motivation: scientific results , **is that enough??**
  - Challenges:
    - Understand application behavior and actual impact (**metrics**)?
    - Limited skills/time/opportunity for optimizations (**transparent**)?
- Admins
  - Motivation: **keep the system running under the limits**
  - Challenges:
    - Cannot support individual users with energy optimization efforts (**automation**)
    - Operate the system within data center power limits (**hard power capping**)
- Data center management
  - Motivation: **maximize work output within CAPEX/OPEX budgets**
  - Challenges:
    - Understand the workloads running in the data center (**analysis**)
    - Control energy budget (**energy optimization**)
    - Charge users for energy consumption (**accounting**)

Stakeholders perspective on Energy Optimization: Main challenges

- Users
  - Some users has expertise on resource optimisation
  - Others not
  - Most of them doesn't have on energy optimization
  - **So what is their benefit for being "energy efficient" ?? → It must be easy!**
- Admins
  - **For them is another tool , so another source of problems → It must be easy!**
- Data center
  - Reduce the electricity consumption and works under the power limits → **It is a must!**

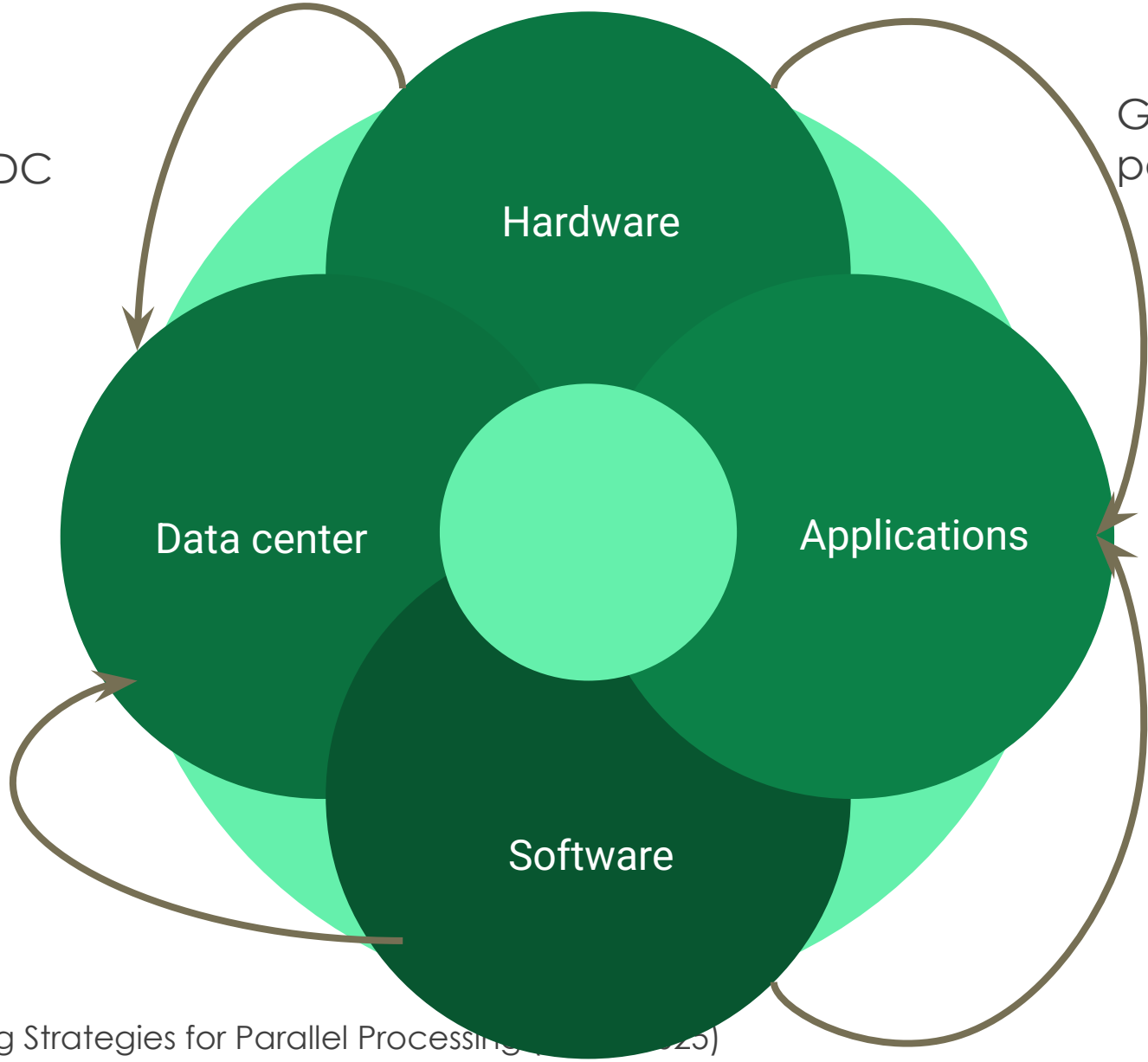
It is pending to see how traditional scheduling policies (CPU hours) are replaced by KWH or Joules or GFLOPS/W targets or ….., energy efficiency metric, and the effect on users

Collaborative work for Energy efficient DC

EAR Energy Aware Runtime

Maximum performance with DC limits

Get maximum performance/watt



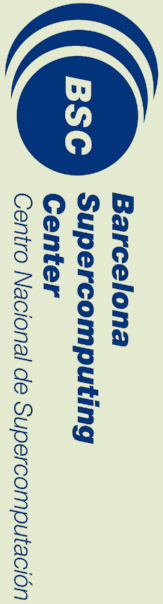
Get the optimal runtime configuration settings

Get the optimal runtime/static configuration settings

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EAR
Energy Aware Runtime

The roadmap to energy efficient Data Centers



28th Workshop on Job Scheduling Strategies for Parallel Processing (JSSPP 2025)

Monitoring

- **Data Center monitoring**
  - Computational and Non-computational devices
  - Different granularities: ms… hours
  - Different sources of data: In-band, out of band, exporters, etc
  - Different scales: few nodes/devices…. Thousand
- **Application monitoring**
  - Different that jobs, fine grained: 1 job multiple applications
  - Metrics : CPU, GPUs, FPGAs, etc
  - Programming languages and frameworks
  - Schedulers
  - Virtualization
  - others…

Optimization

- Applications and system can be dynamically configured to adapt to the running ***workload***
- **What is a workload**? It depends on the context
  - Run vs Idle period → Select the best CPU governor. AFAIK there is nothing similar on the GPU
  - Running an application → Select the best CPU/Memory/GPU frequency (and governor) resulting in the maximum performance/watt
  - Running a workflow → Select the best per-application configuration taking into account the whole workflow
  - DC workload → Update the system policy (and policy arguments) to maximize DC goals or constraints

Power/thermal/Cooling management

- Optimization can be seen as a ***ratio*** , power, thermal, cooling management is about infrastructure limits
  - Limits than cannot be exceeded
- This technology has been historically provided by Vendors since it is a “hardware” issue
- Several strategies
  - Preventing problems → limit always set by hardware
  - Predicting/Detecting → There is a risk
  - Dynamic settings → Adapting settings to the workload
  - ***Combining Prediction & Dynamic management***

Analythics

- Since many years some Data Centers have shared their workload traces for job scheduling research and evaluation
- **This data was used mostly with scheduling policies research purposes**
  - **No application performance**
  - **No malleability**
  - **No “power” related metrics**
- Data with individual application characterization
- Energy and power models has been created using data from applications

Analythics (II)

- There is not yet an accepted extension of SWF
  - We proposed Modular workload format few years ago in this workshop
- **Surprisingly…. No new traces, no new models but still strategic!!**
- In some institutions there is more access to collected data, lots of groups doing research now in particular with AI/ML technologies
  - Power prediction
  - Anomaly detection
  - etc

EAR
Energy Aware Runtime

The EAR roadmap to energy efficient Data Centers

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

28th Workshop on Job Scheduling Strategies for Parallel Processing (JSSPP 2025)

EAR and the Stakeholders

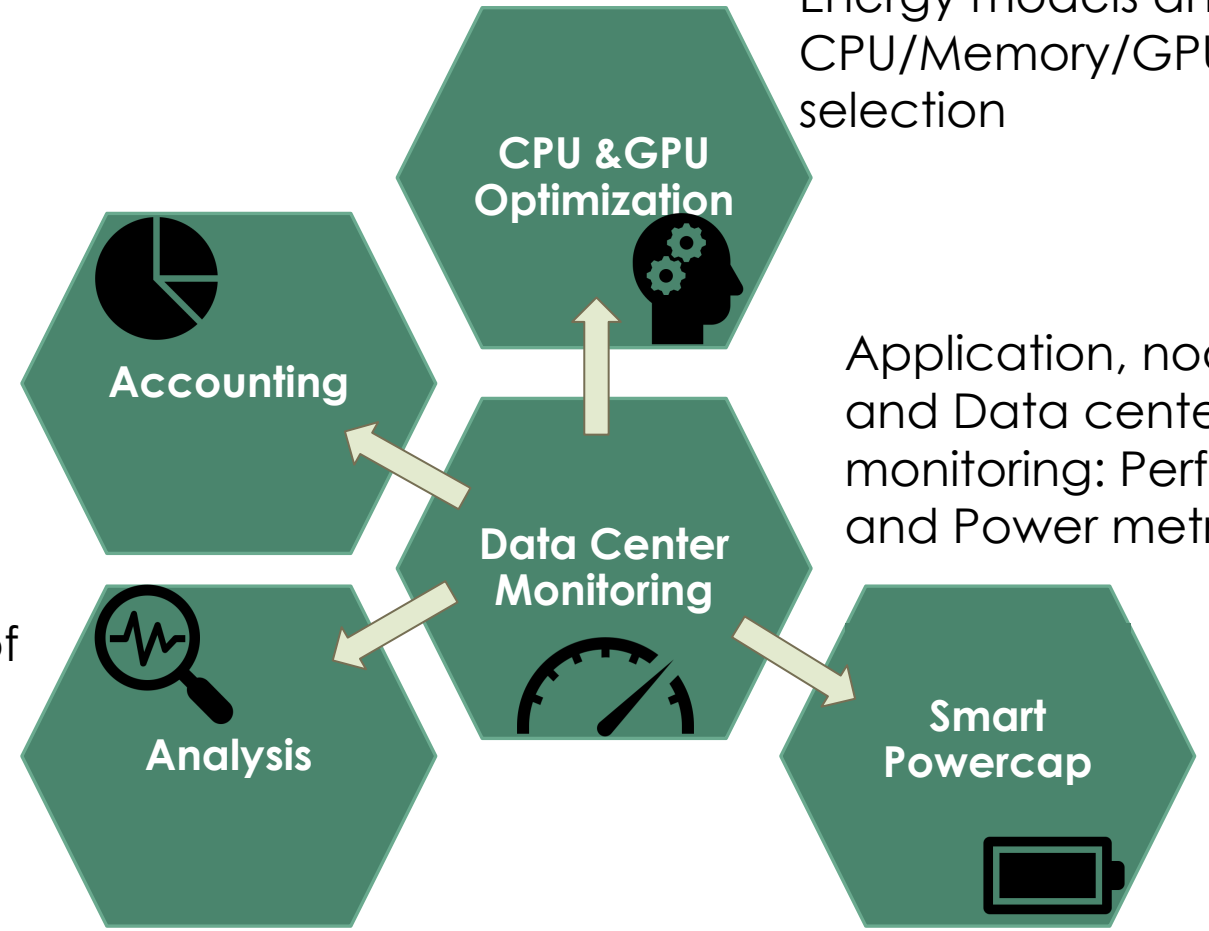
- Users
  - **Automatic (easy)**
  - **Portable**, it's worth to do the effort of using EAR tools
  - HPC and AI frameworks, programming models etc supported
- Admins
  - It's a new tool. Yes, but very **similar to existing tools** (SLURM)
  - Very few dependencies
  - Powerful deployment and support for heterogeneous systems
- Data center management
  - Power management
  - Energy optimization
  - Non-computational devices(WIP)

EAR : System software for Energy efficiency

EAR Energy Aware Runtime

Energy models and policies for CPU/Memory/GPU frequency selection

Data reporting for accounting, billing, visualization and analytics



Application, node , cluster and Data center monitoring: Performance and Power metrics

Analysis and classification of application metrics for energy and power optimization

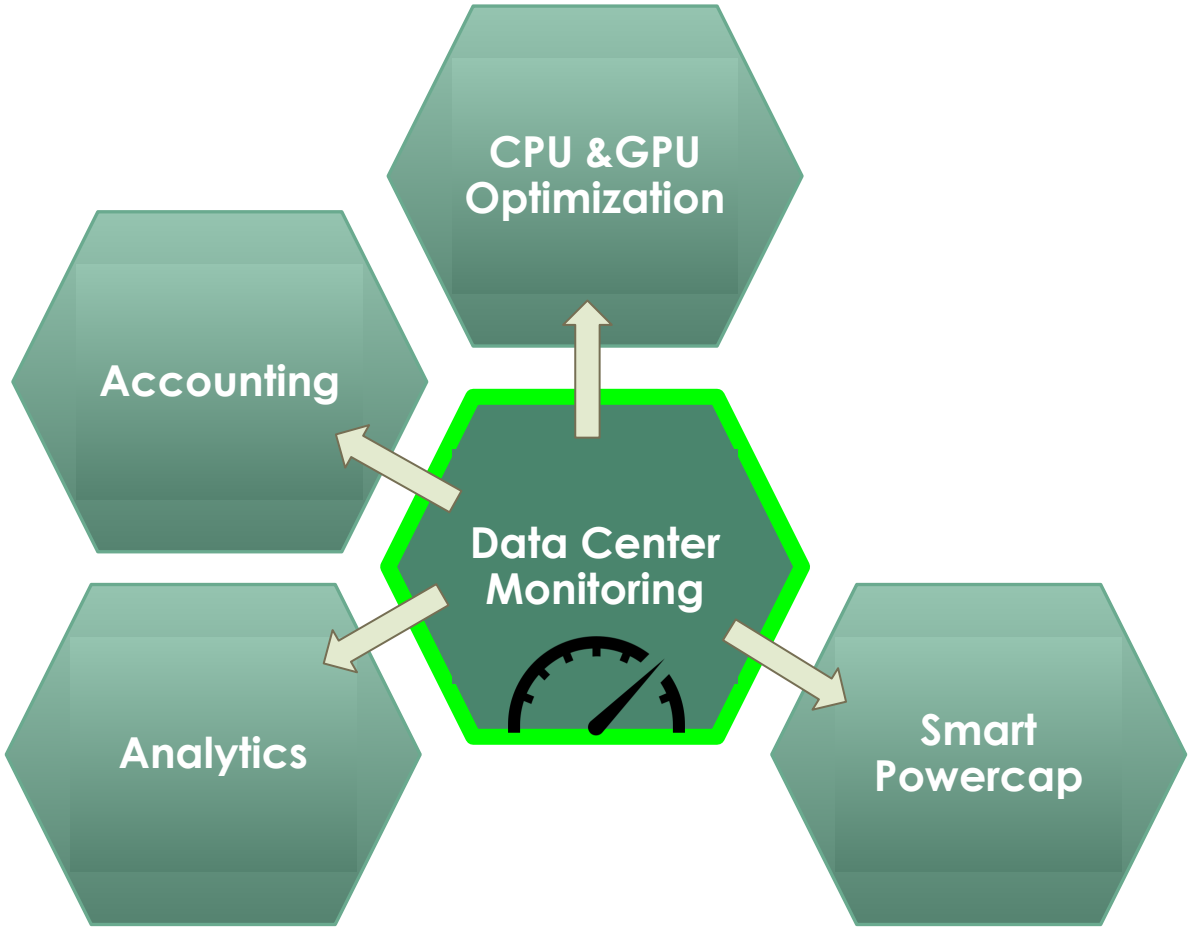
Power control to guarantee data center operational limits

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Data Center Monitoring and Optimization

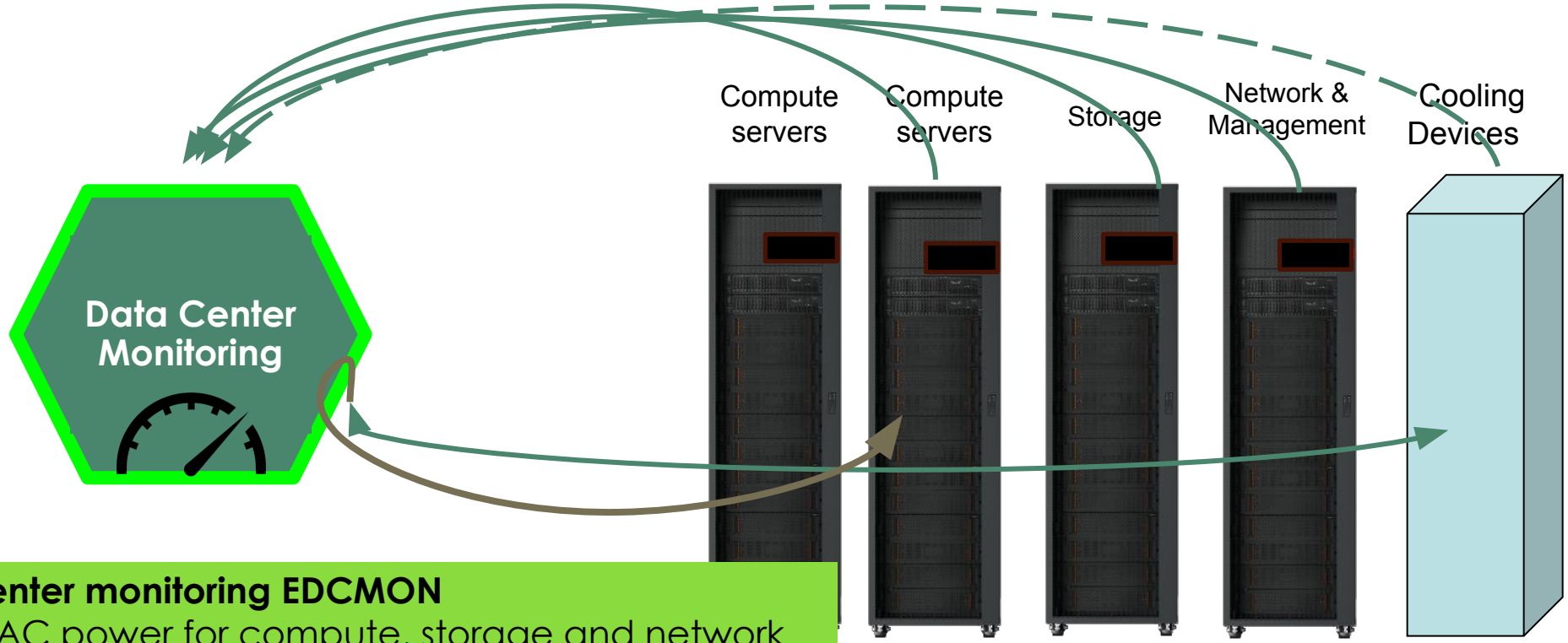
- Monitoring as source for
  - Optimizing
  - Management

- Compatible with other tools
  - Sharing data
  - European software stack



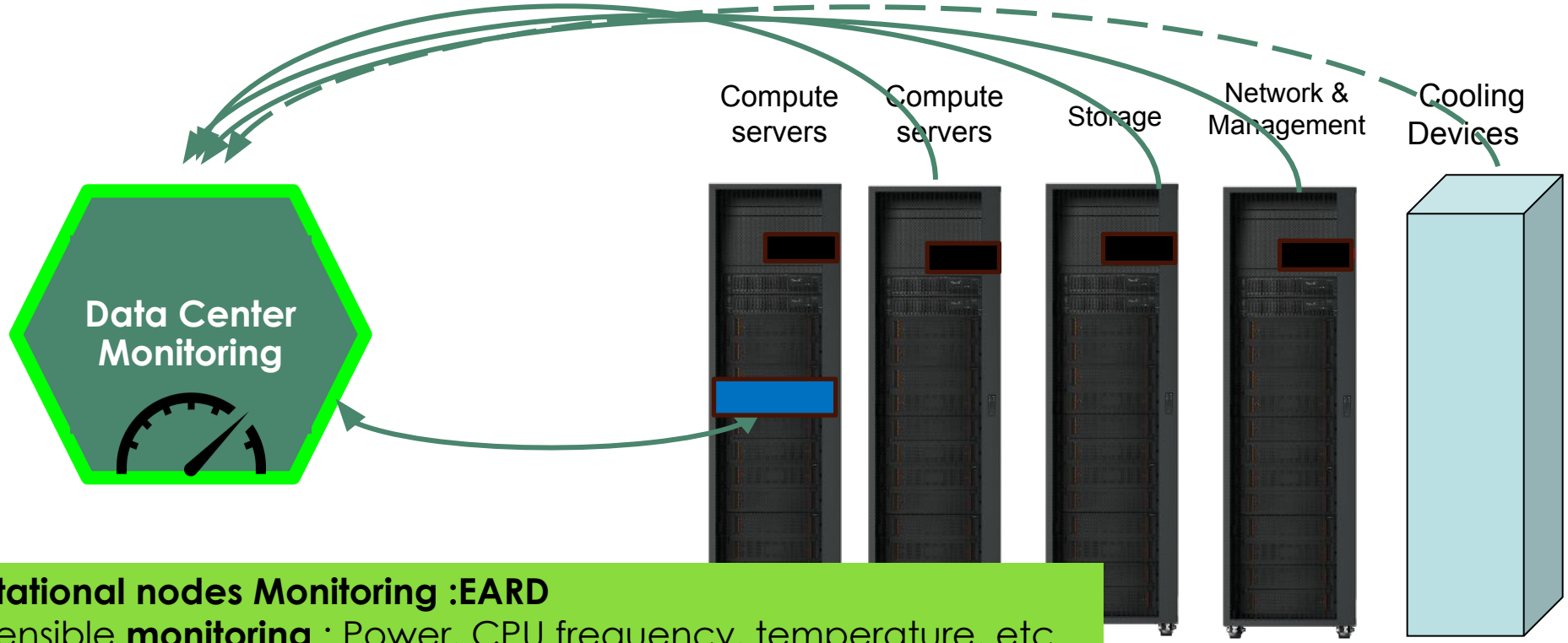
REGALE project

Rack, Node, Application Monitoring



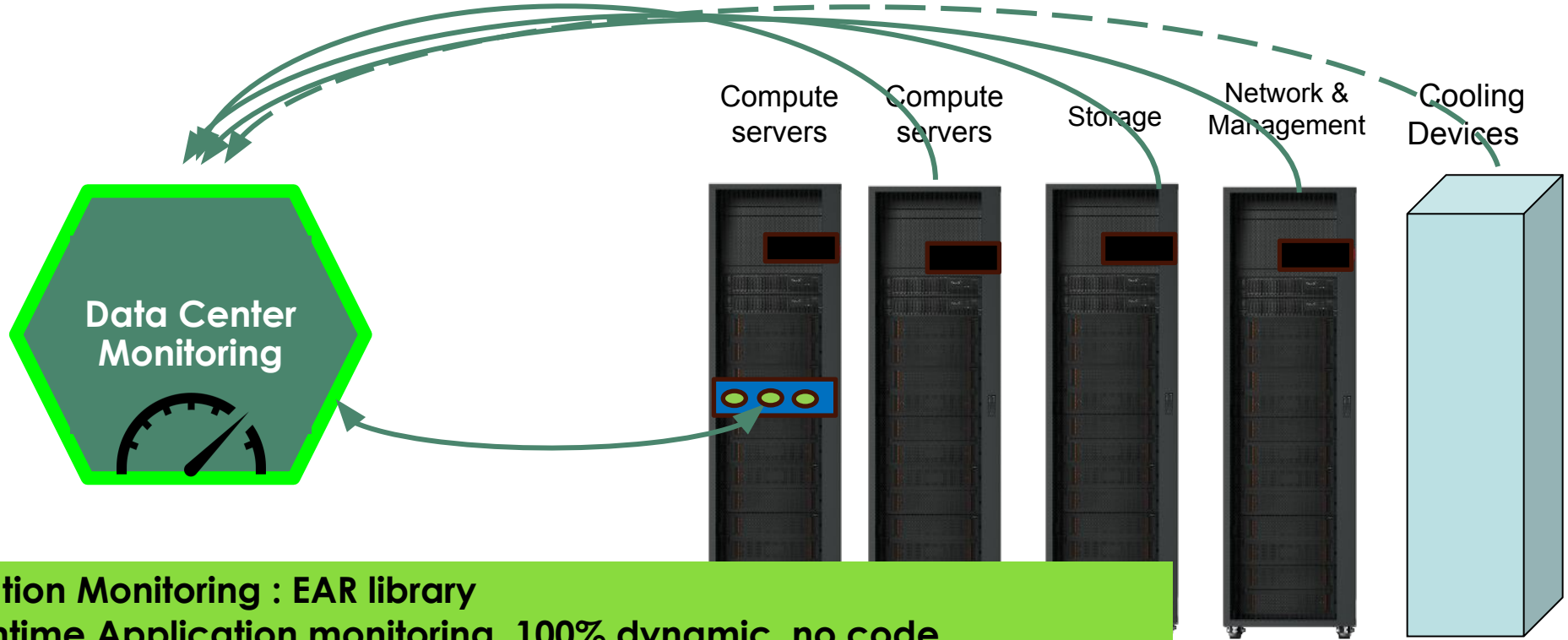
- **Data Center monitoring EDCMON**
  - AC power for compute, storage and network through the PDUs
  - Report to DB
  - Visualization
  - Integration with EAR powercap service(WIP)

Rack, **Node**, Application Monitoring



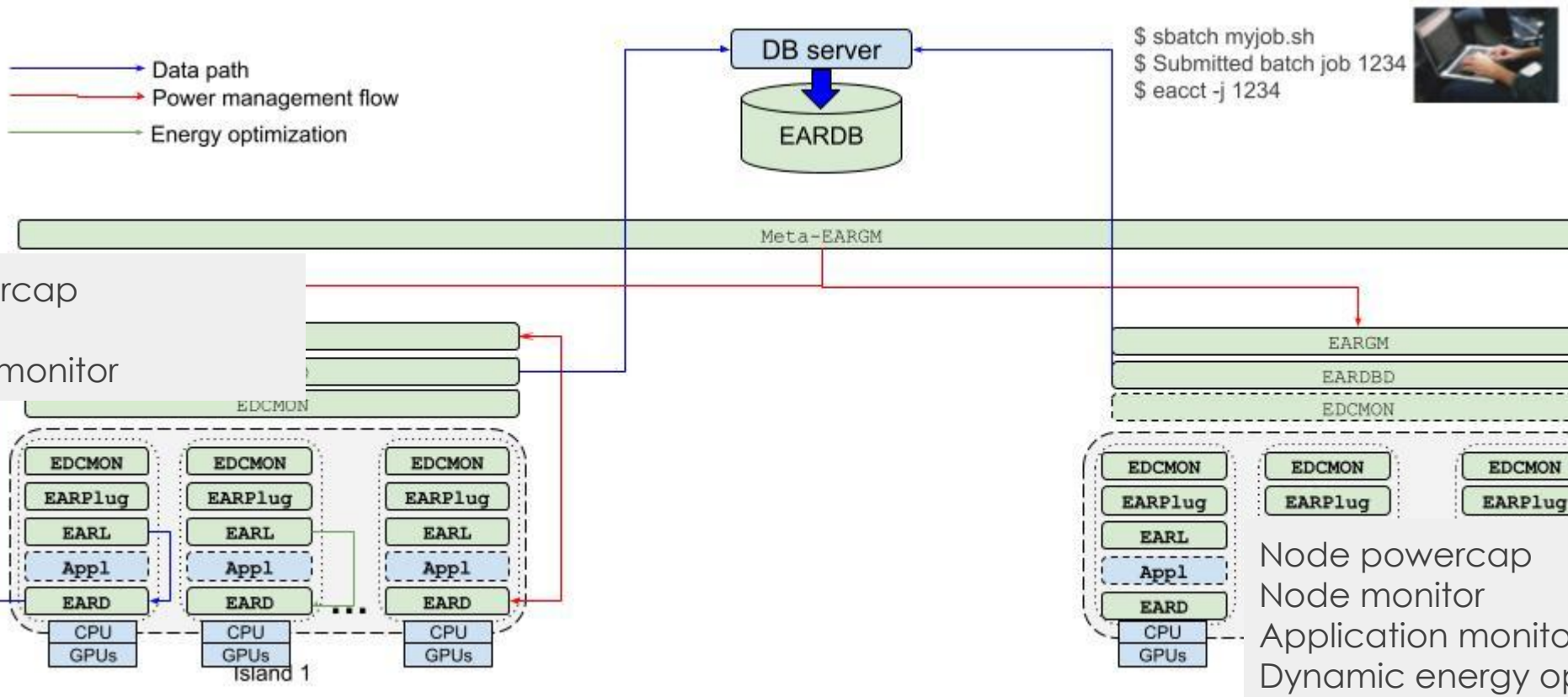
- **Computational nodes Monitoring :EARD**
  - Extensible **monitoring** : Power, CPU frequency, temperature, etc
  - Multiple sources of data: inband IPMI, redfish, NVIDIA, MSR, RAPL…
  - Extensible **report** : MariaDB, Postgres, Sysfs, Prometheus,…
  - **Anomalies** detection for power and temperature

Rack, Node, **Application** Monitoring



- **Application Monitoring : EAR library**
  - **Runtime Application monitoring, 100% dynamic, no code modifications**
  - **Collects the Application Runtime Signature**: Performance & Power

EAR architecture



Cluster powercap
DB report
Data center monitor

Node powercap
Node monitor
Application monitoring
Dynamic energy optimization

- Software group of nodes
- Distributed features: Monitoring, optimization, power management, reporting
- Vendor independent

EAR
Energy Aware Runtime

Dynamic energy optimization

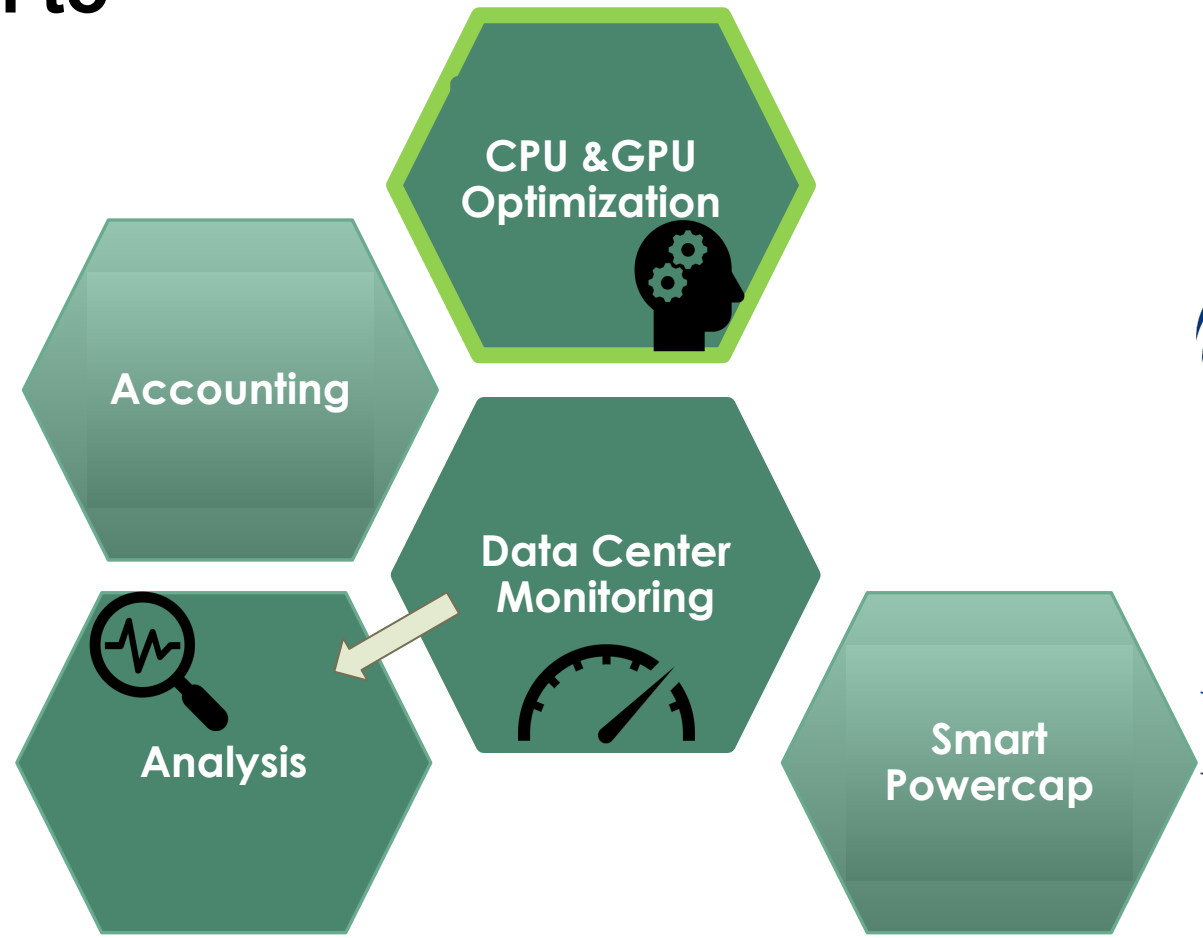
Get more performance per Watt

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

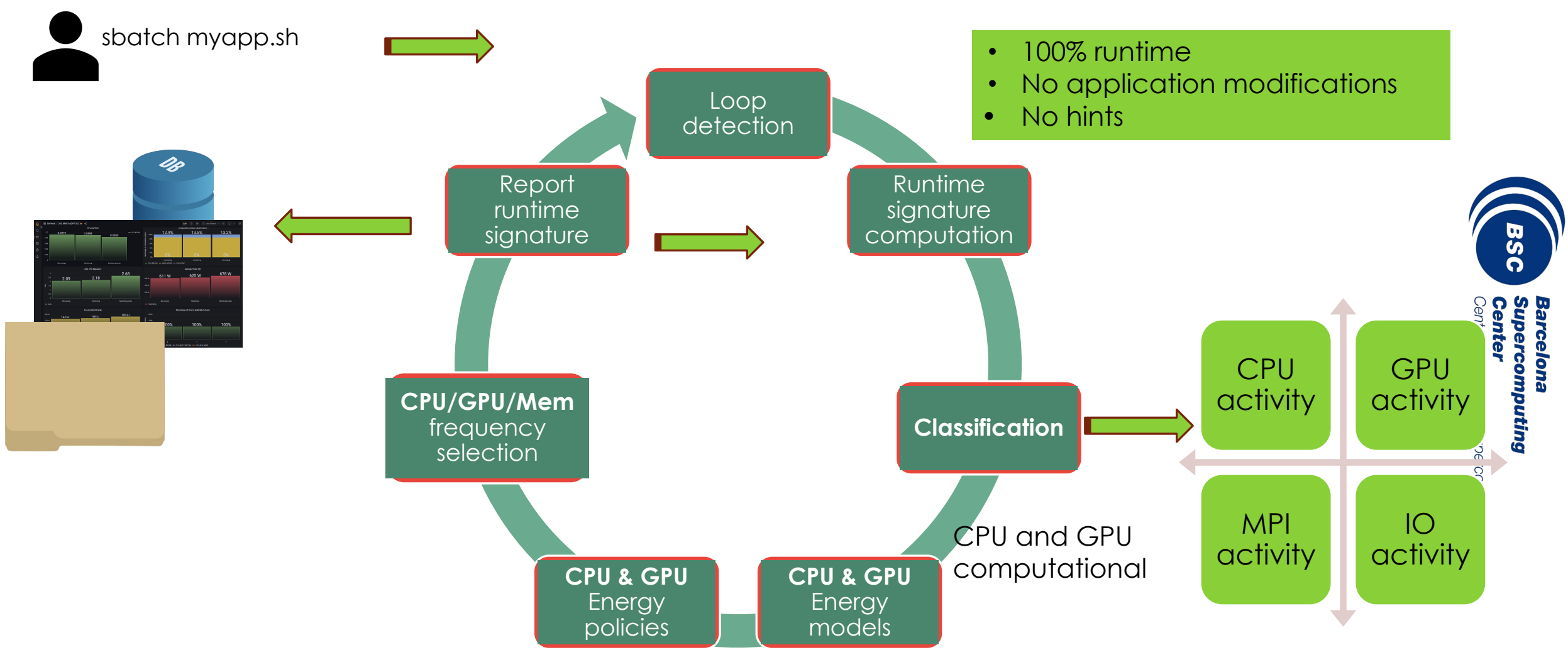
Dynamic Energy Optimization: Adapt the power consumption to the application activity

- Runtime
- Transparent
- Application : Adapt configuration to **application** “Activity”
- Cluster: Adapt configuration to **Workload** “Activity”

- **Application high efficiency mind-set**

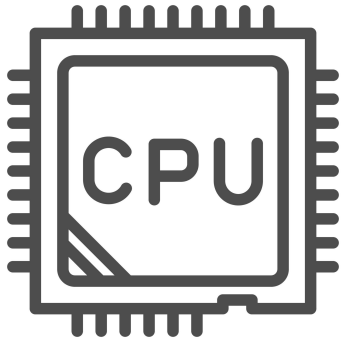


The EAR Library optimization loop

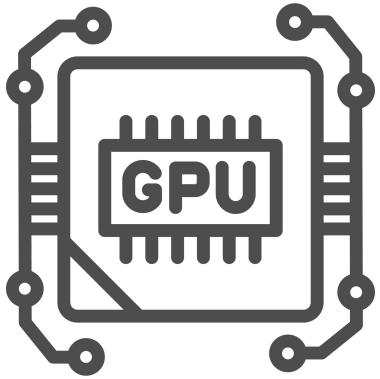


Application runtime characterization: Utilization vs “effective” utilization: The signature

The key question is not whether the CPU or GPU is *used,* is **what is used for and how much does it consumes** to do the job…. And EAR knows!!!



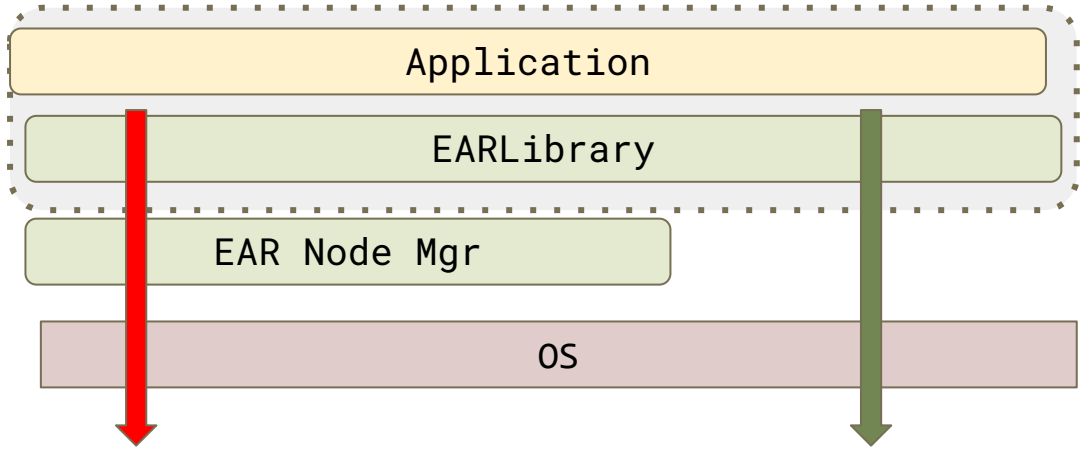
CPU activity : CPI, Memory BW, GFlops, IO, network
Power consumption: DC node power, DRAM and CPU power



GPU activity : GPU and memory utilization, GPU flops (NVIDIA)
Power consumption: DC node power, GPU power

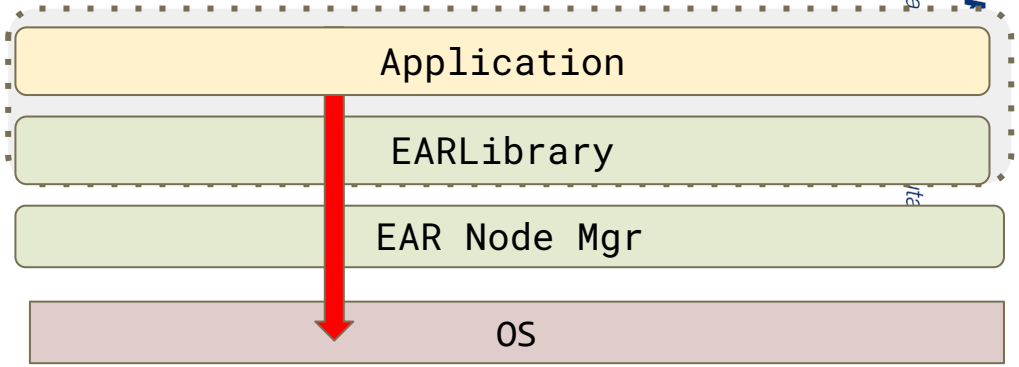
What about portability?

- EAR is very portable and flexible since all the APIs for monitoring and management are implemented in **two layers**
  - Generic
  - Device/API specific
- There is always a **dummy API** in case one API is not supported for a new architecture



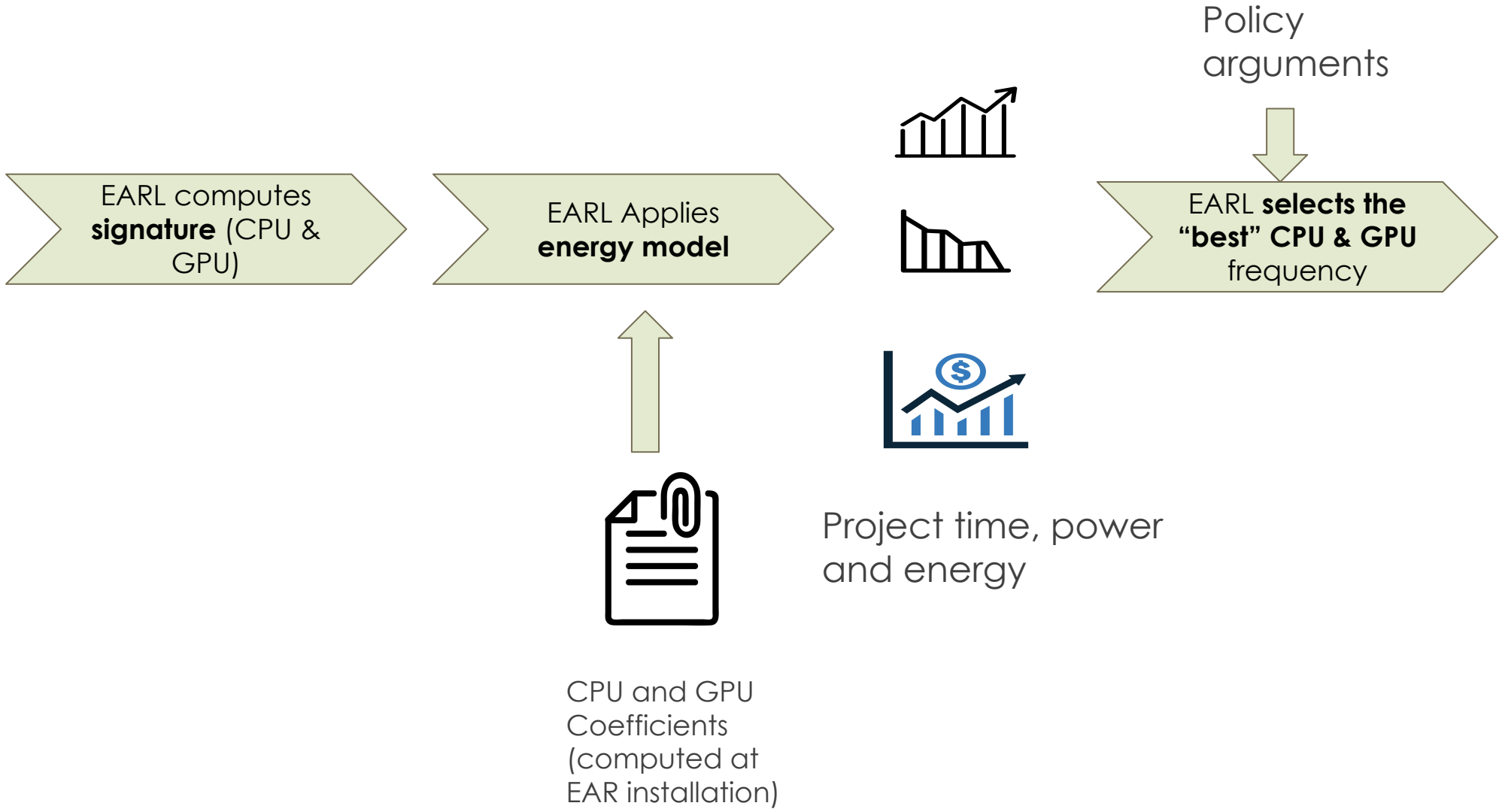
IPMI RAPL MSR HSMP REDFISH etc

Perf nvidia levelzero rsmi hwmon etc



- ➔ OS Cpufreq : userspace, intel_pstate
- ➔ AMD HSMP with virtual list of p-states
- ➔ GPU
  - Libnvdidia-ml, Levelzero, smi

Energy dynamic optimization



Min_energy_to_solution policy (CPU and **GPU**): Minimizes the **energy with a maximum performance (time) penalty defined as threshold**

Results on H100 - SXM5

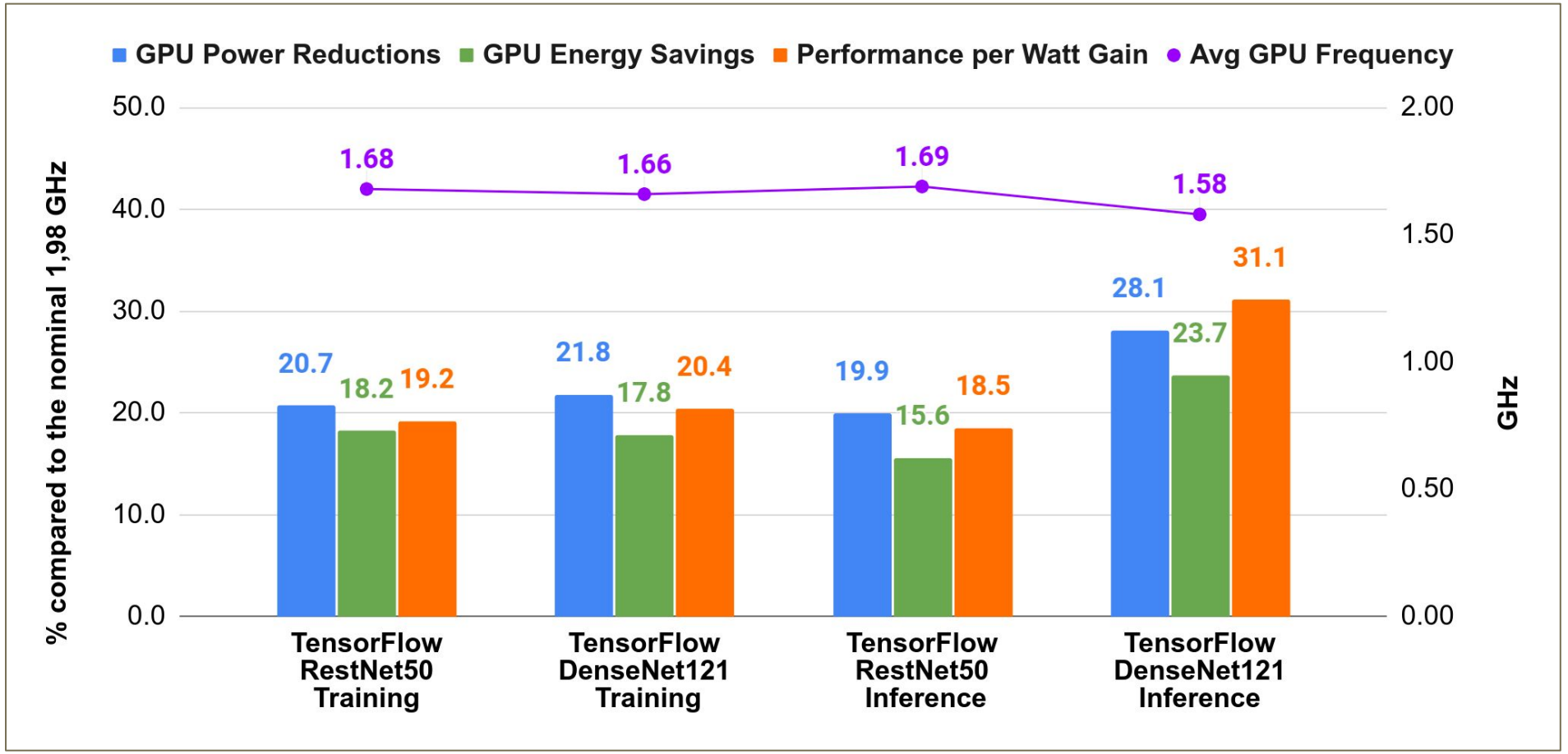
- 4x GPU
- TDP: 700 W
- Nominal Frequency: 1980 MHz

AI Applications on H100

Applications	Version	Model	# GPUS	MPI tasks	OpenMP Threads	Inputs (batch_size)
TensorFlow	2.15	Resnet 50	1	1	72	128
		DenseNet121	1	1	72	128
		VGG19	1	1	72	128
PyTorch	2.2	Resnet 50	1	1	8	256
		Resnet 50	4	4	4	256
		DenseNet121	1	1	8	128
		DenseNet121	4	4	4	128

EAR GPU optimization on AI applications (1 GPU)

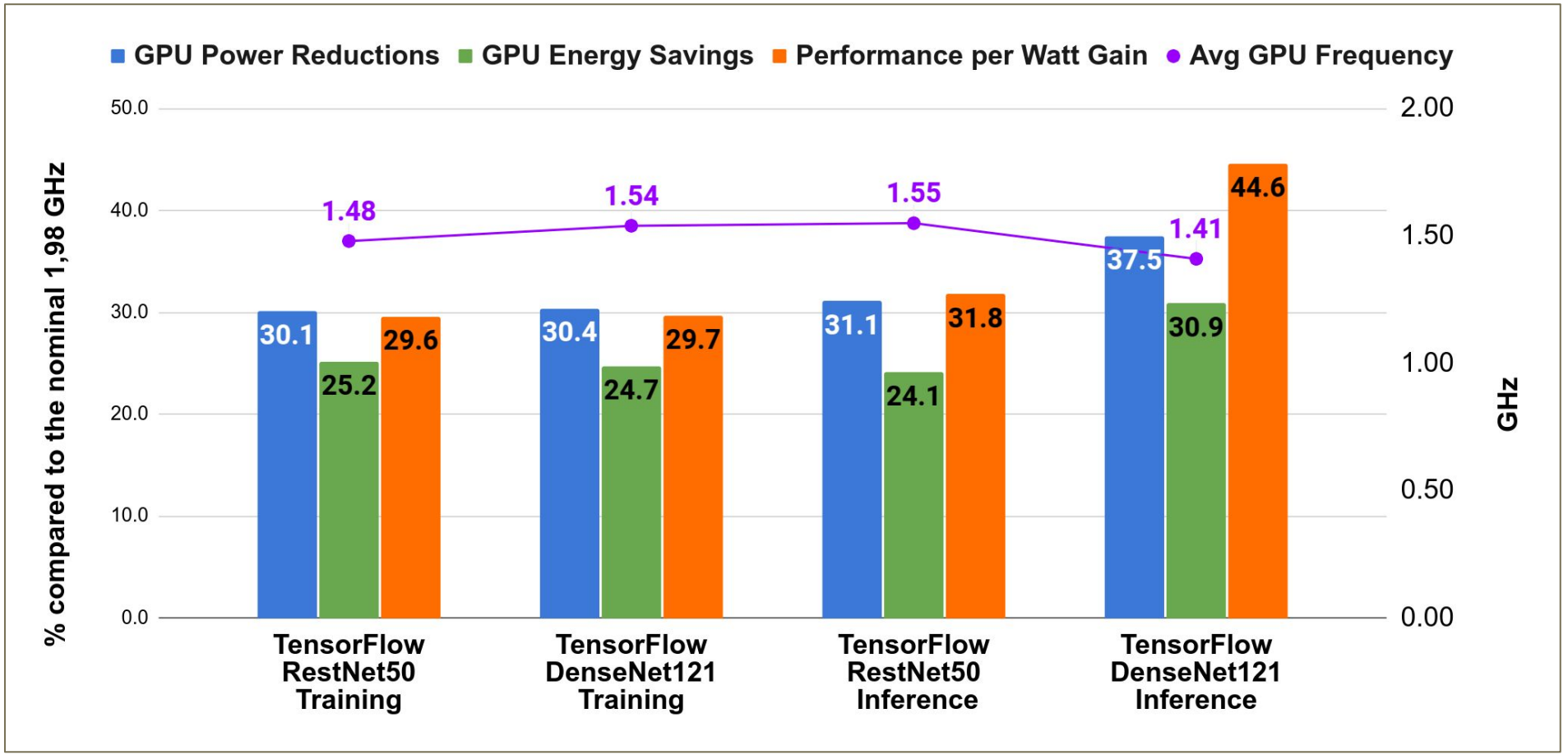
TensorFlow Benchmarks: GPU Energy savings and efficiency



EAR Optimization with 5% policythreshold: **Avg energy saving 18.8 %**

EAR GPU optimization on AI applications (1 GPU)

TensorFlow Benchmarks: GPU Energy savings and efficiency

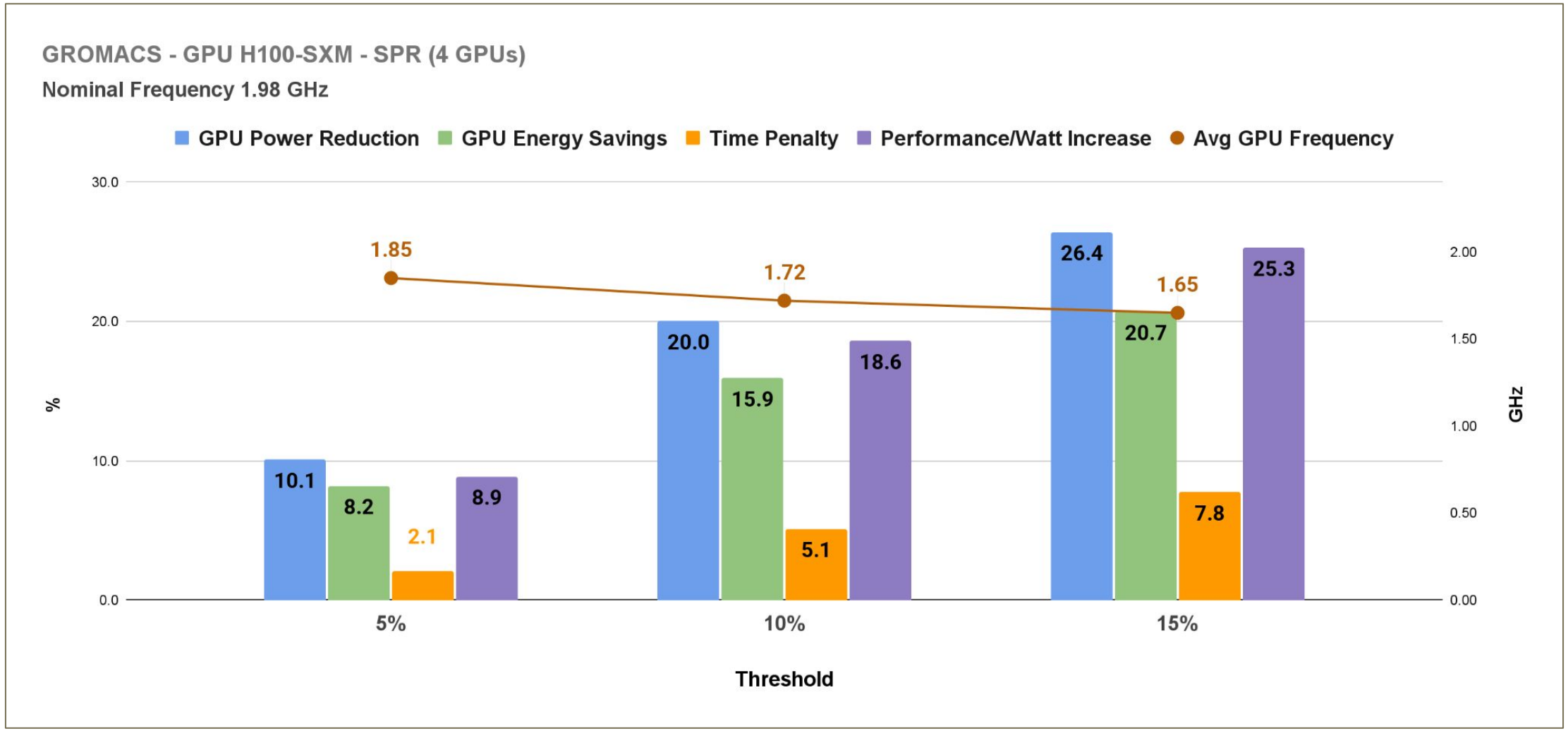


EAR Optimization with 10% threshold: **avg. energy saving 26.2%**

Gromacs on H100 SXM5 on 4 GPUs

- Processor: Xeon Platinum 8462Y+ (S:C:T = 2:32:2)
- GPU: H100 SXM5, TDP = 700 W/GPU
- Gromacs dataset: BenchPEP-h
- GPU min_energy policy:
- Benchmark configuration:
  - 4 GPUs with 4 MPI tasks and 16 OpenMP threads

GPU savings with EAR Optimization



EAR
Energy Aware Runtime

Smart powercap

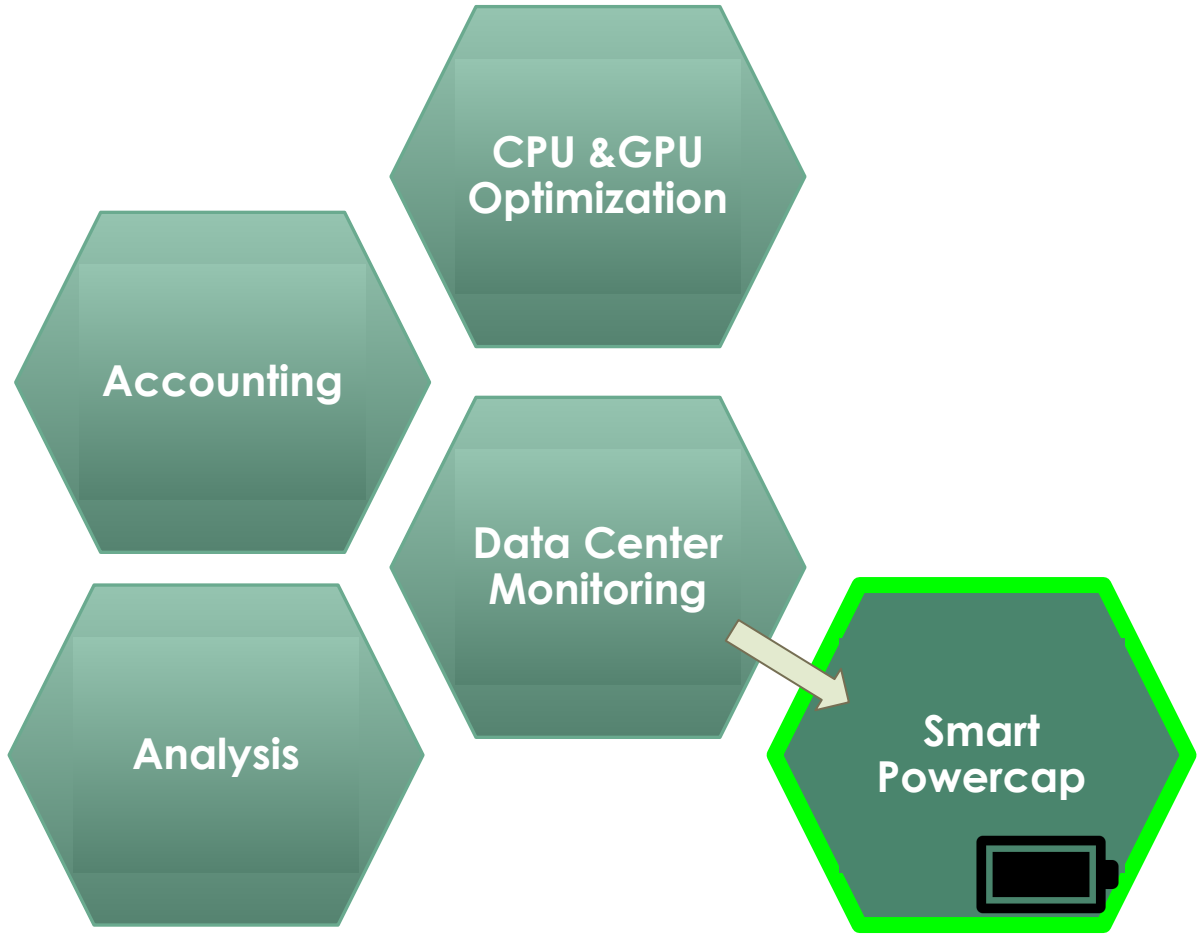
Dynamic powercap guided by application activity

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

28th Workshop on Job Scheduling Strategies for Parallel Processing (JSSPP 2025)

Smart Powercap

- Node powercap
  - Heterogeneous
  - Dynamic
  - Application “driven”
- Cluster powercap
  - Heterogeneous
  - Dynamic
  - Node (Application) “driven”
- **Cluster high efficiency mind-set**

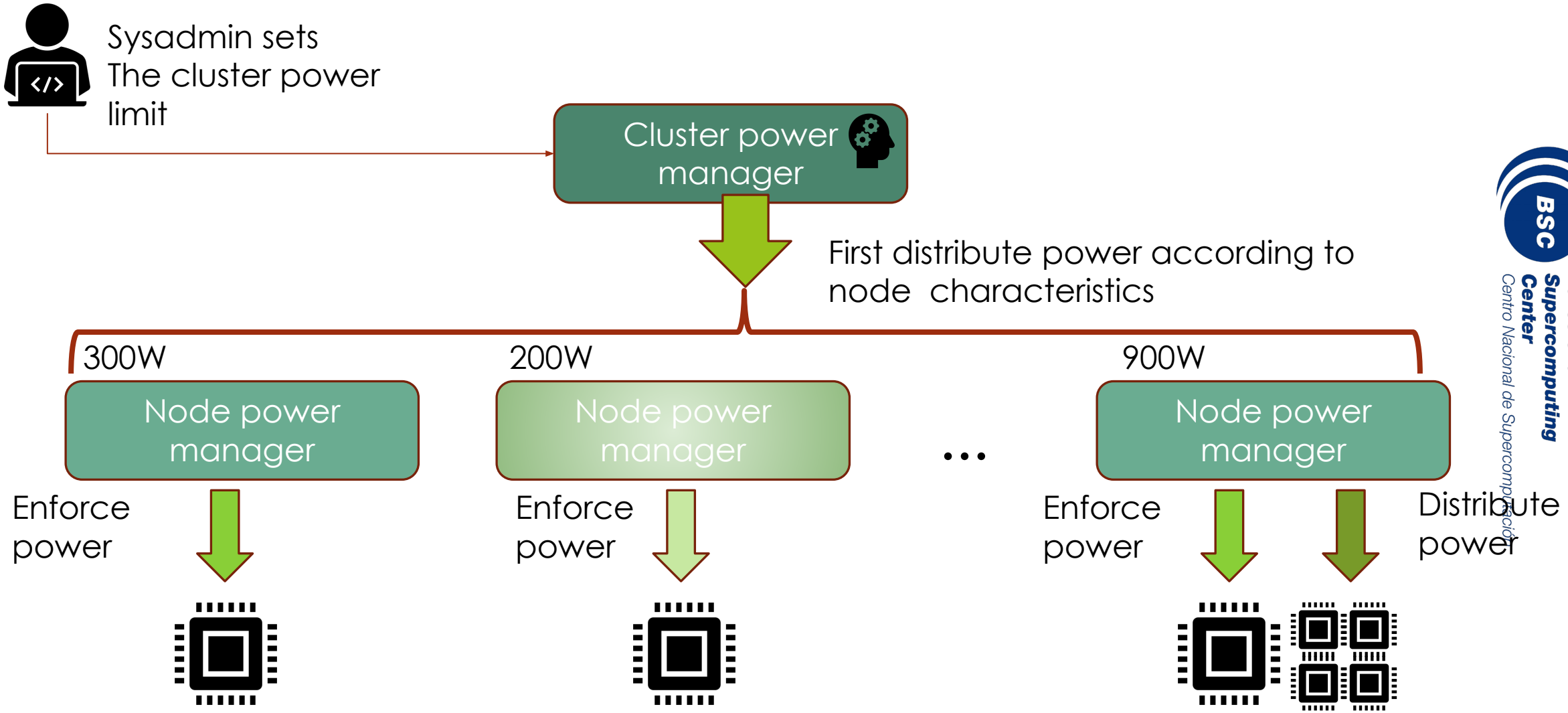


EAR smart powercap overview

- **Cluster power manager** distributes power to computational nodes
  - Hierarchical architecture for large scale clusters: 1 Meta- EARGM, N EARGMs
  - Two algorithms offered: **soft** (lightweight) and **hard** powercap
- **EAR Node powercap manager** enforces node power limit
  - Extensible through plugins: CPU, GPU
  - Dynamic **intra-node power re-allocation** based on application activity
- **EAR library informs** the EAR node powercap manager of application activity and power requirements
- Architecture
  - Software architecture: Compute nodes can be "grouped" with flexible & dynamic criteria
  - Vendor independent

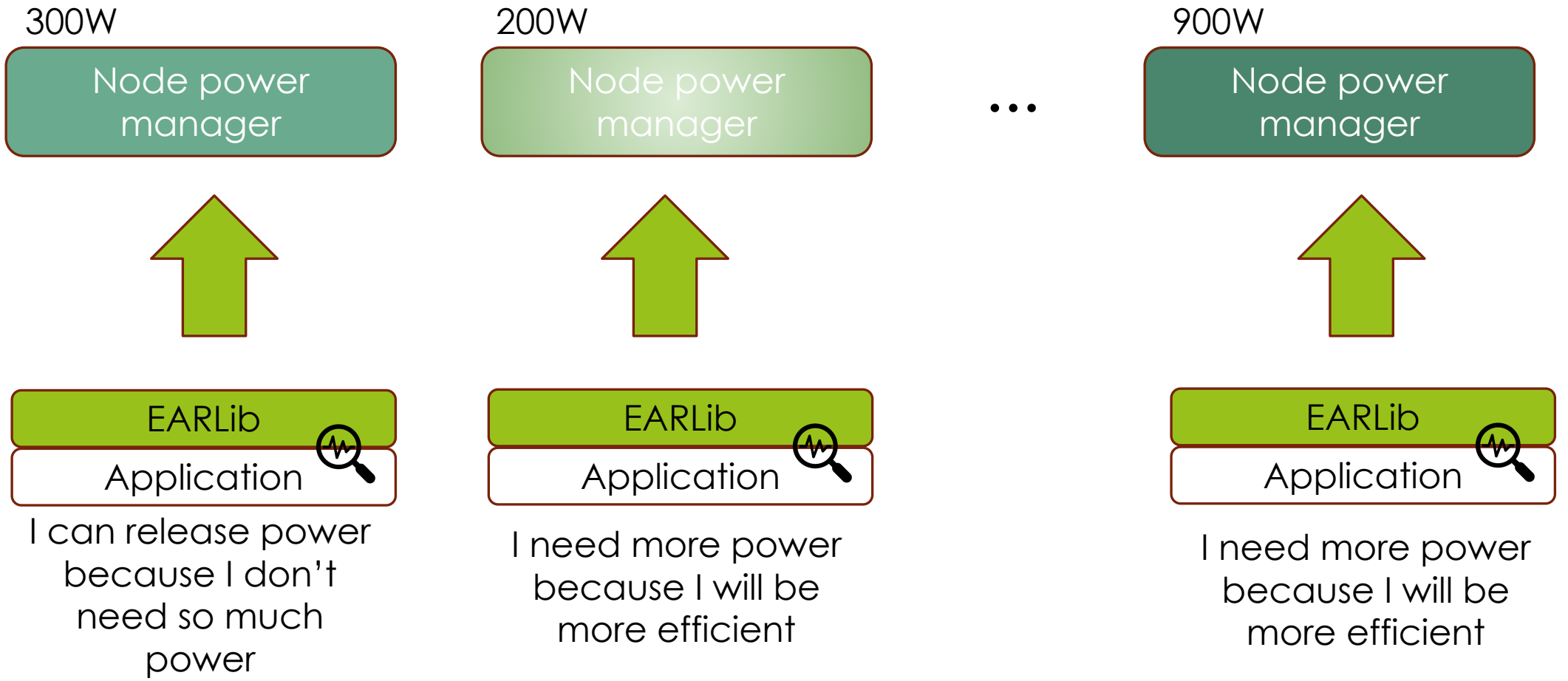
Powercap (I): Initial distribution

Cluster power manager distributes power and node power manager enforces power

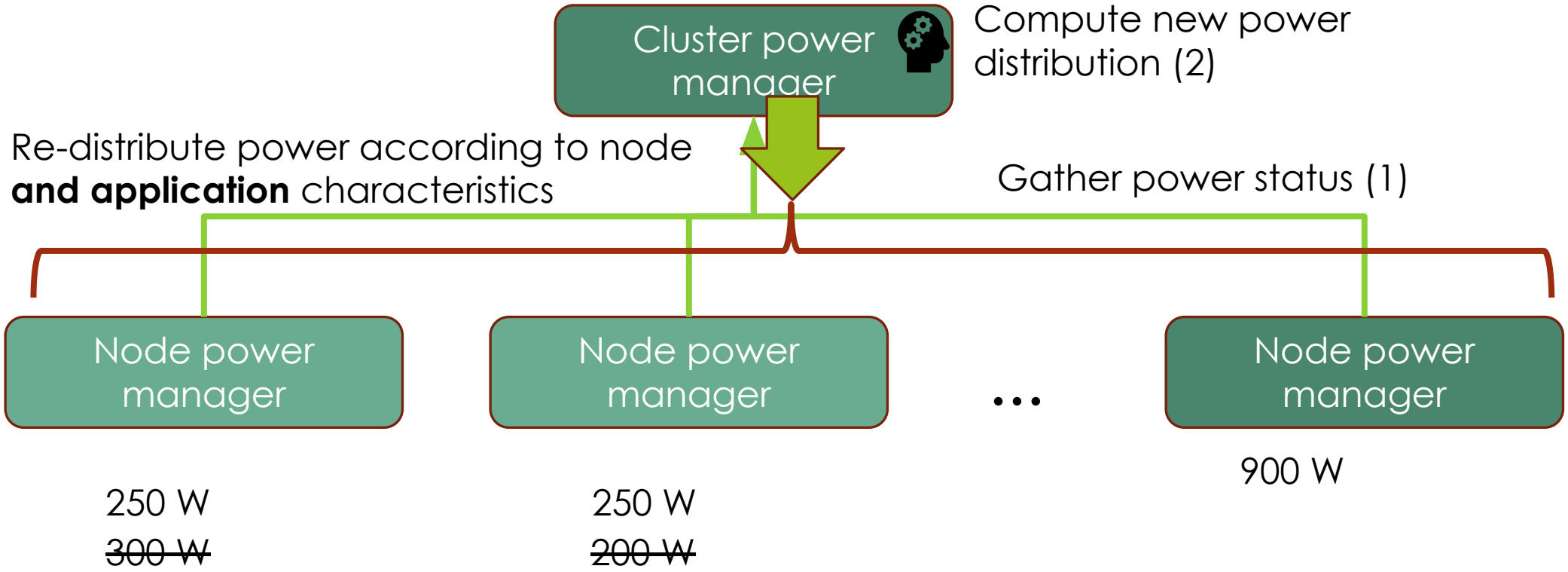


Powercap(II): Application feedback

Application (through EARlib) informs each node about its power needs



Powercap(III): Dynamic power reallocation



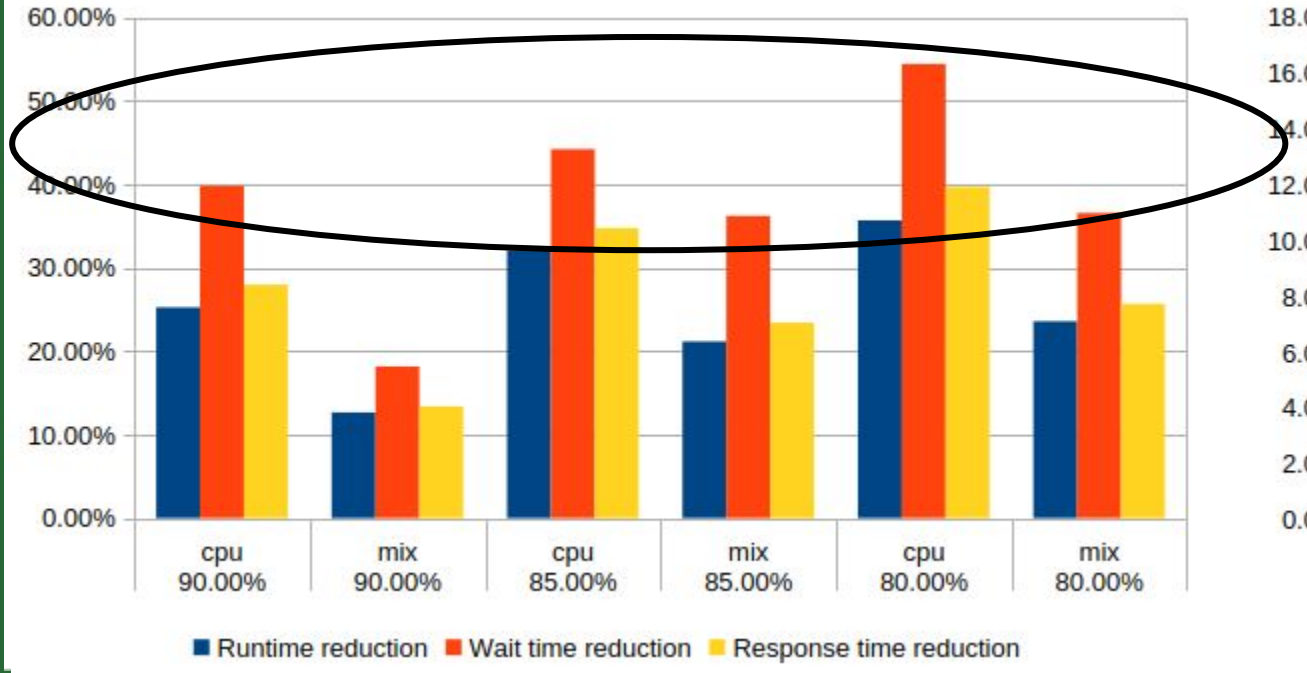
The “High Efficient Computing” mind-set

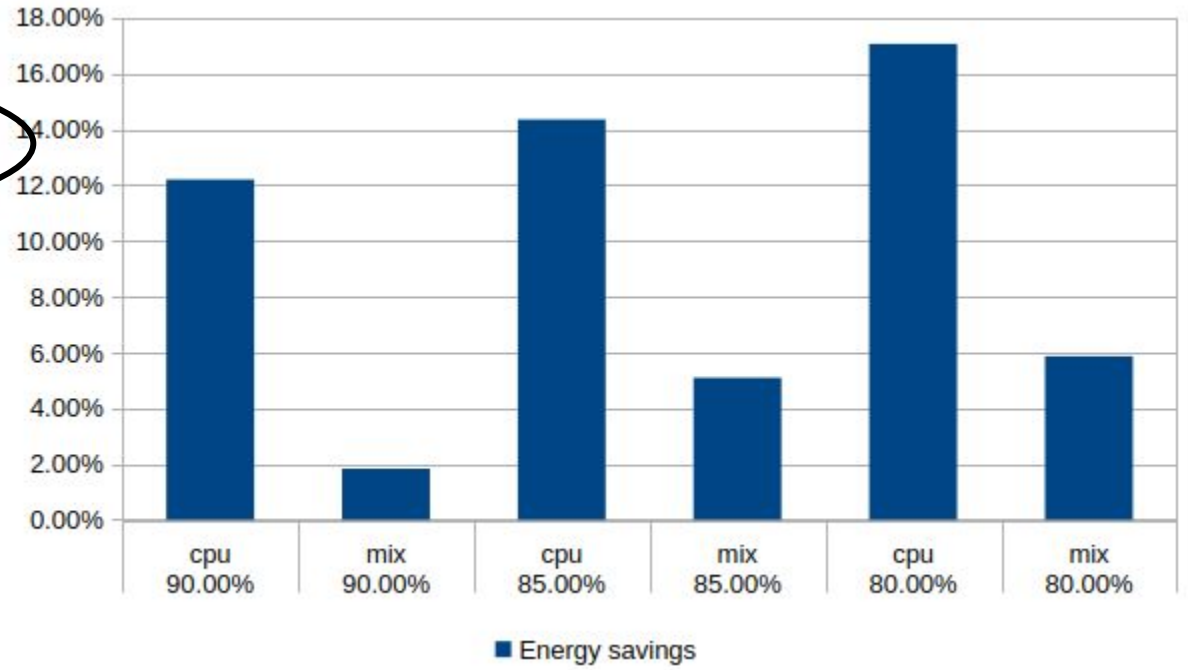
- Energy optimization seems to be against maximizing performance , but basically we look at individual applications and unlimited resources

- **This is not longer true if we focus on workloads and limited systems**

Energy/Power optimization is not a pain , it is a MUST

EAR cluster hard-powercap vs static powercap





- 10K nodes (simulation based results)
- SPR characterization CPU nodes

EAR
Energy Aware Runtime

Reporting and visualization

- EAR CLI
- EAR visualization tools
- Grafana

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EAR command for job accounting and node monitoring: eacct and ereport

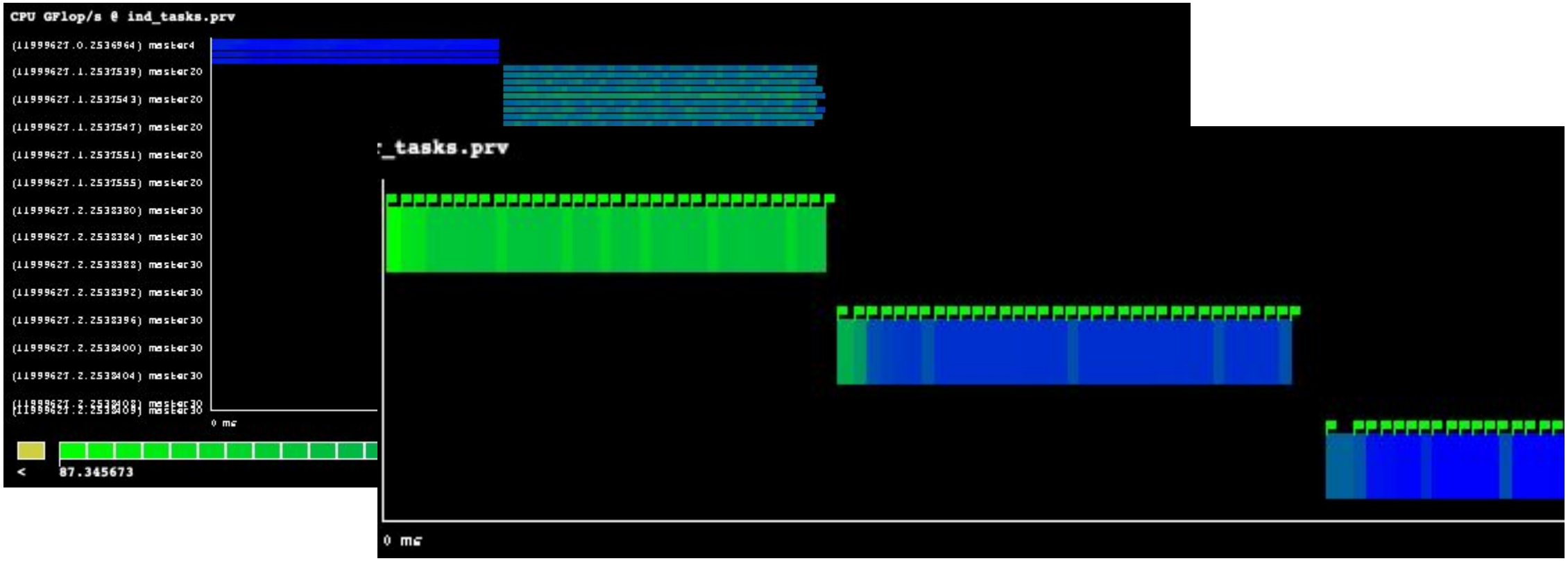
- eacct is a linux command line tool to gather job metrics
- Regular users can gather their own data and managers can gather data from all the users
- Multiple filters and flags supported , including saving all the data in csv files (runtime and application signatures

```
[julitac@int3 TENSORFLOW_TEST]$ eacct -j 1460643
JOB-STEP USER     APPLICATION  POLICY NODES AVG/DEF/IMC(GHz) TIME POWER(W) GBS CPI  ENERGY(J)   GFLOPS/W IO(MBs) MPI% G-POW  G-FREQ  G-UTIL
1460643-5  julitac  rfm_TensorFlowSA MT     8  2.86/2.20/2.12  376.57  1343.82  21.43  0.29  4048335     0.0000  1960.8  44.6  868.93 1.410   89%/46%
1460643-4  julitac  rfm_TensorFlowSA ME     8  2.38/2.40/2.19  374.49  1321.24  21.71  0.30  3958278     0.0000  1971.8  46.7  870.58 1.410   90%/47%
1460643-3  julitac  rfm_TensorFlowSA MO     8  2.38/2.40/2.19  374.28  1322.08  21.90  0.30  3958678     0.0000  1972.7  46.2  869.99 1.410   90%/47%
1460643-2  julitac  rfm_TensorFlowSA MT     8  2.86/2.20/2.10  383.57  1329.94  21.08  0.29  4080972     0.0000  1924.8  44.5  856.43 1.410   88%/45%
1460643-1  julitac  rfm_TensorFlowSA ME     8  2.38/2.40/2.19  374.41  1321.40  21.77  0.29  3957900     0.0000  1972.1  46.2  871.06 1.410   90%/47%
1460643-0  julitac  rfm_TensorFlowSA MO     8  2.38/2.40/2.19  377.26  1313.15  21.66  0.29  3963175  0.0000  1957.1  46.4  863.03 1.410   90%/46%
```

EAR job visualizer

Helps the end users to understand their application behavior and their resources usage.

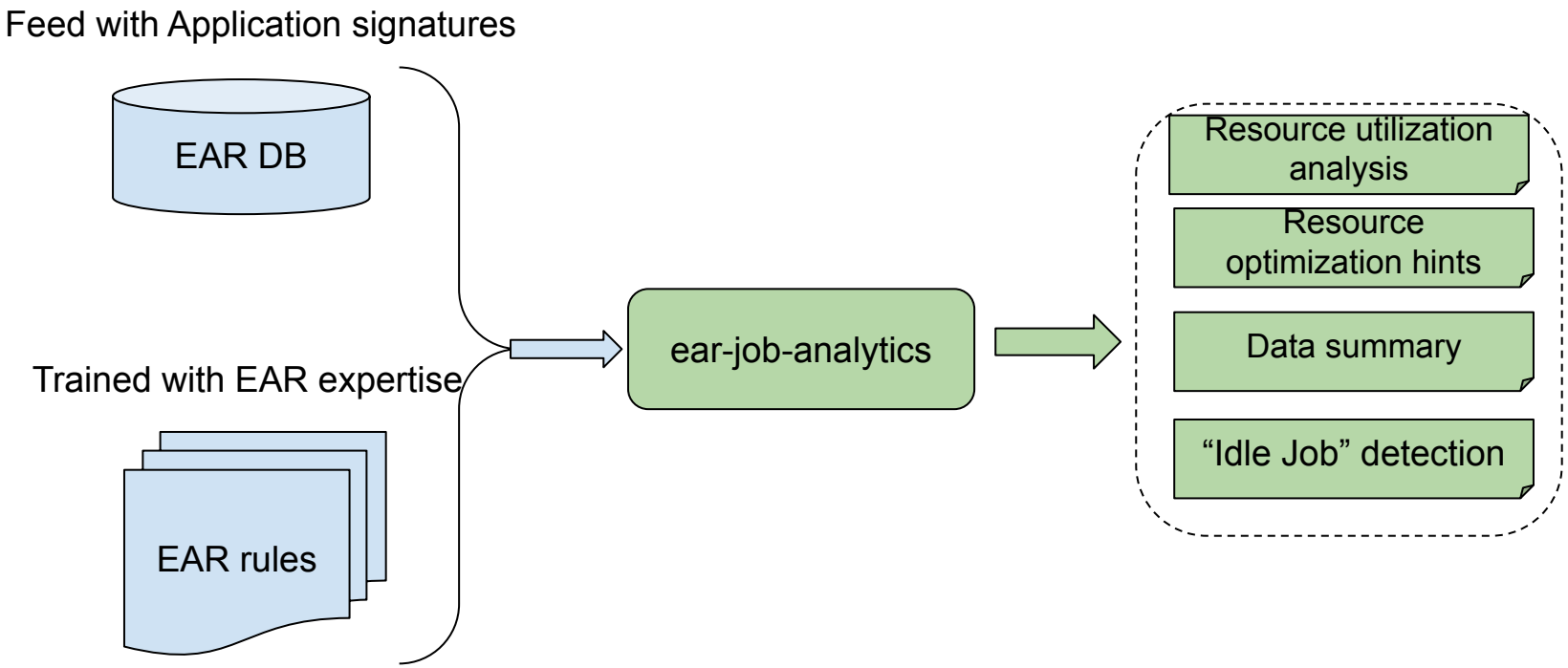
- Static images/Paraver traces with EAR metrics visualization for specific jobs over the time.



Application and system Analytics

EAR monitors jobs and system at runtime → We improve the system utilization with EAR knowledge!

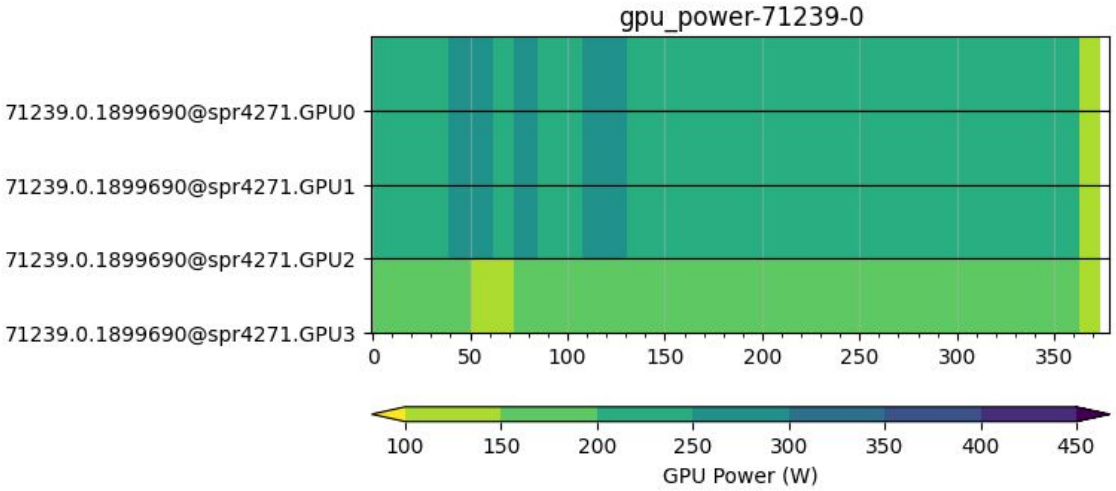
EAR job analytics design

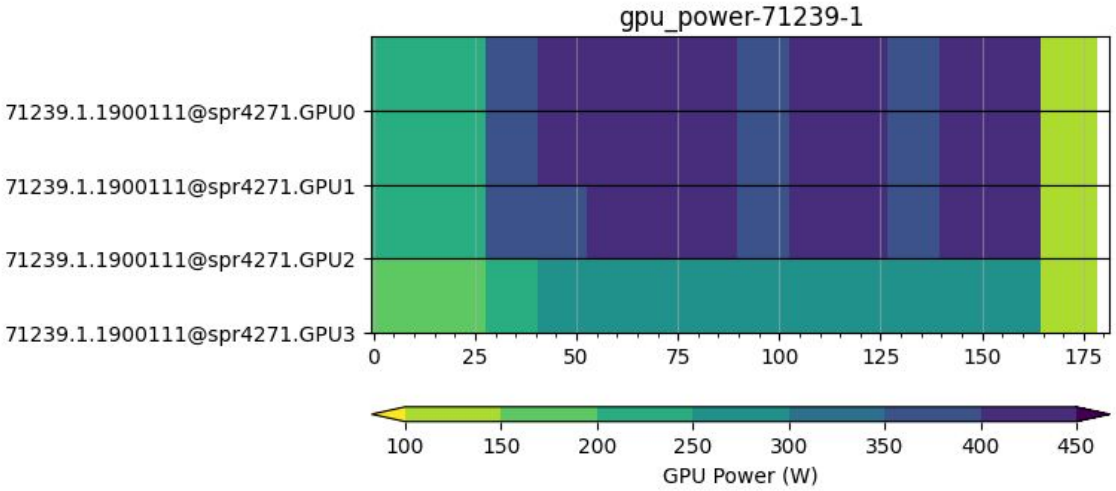


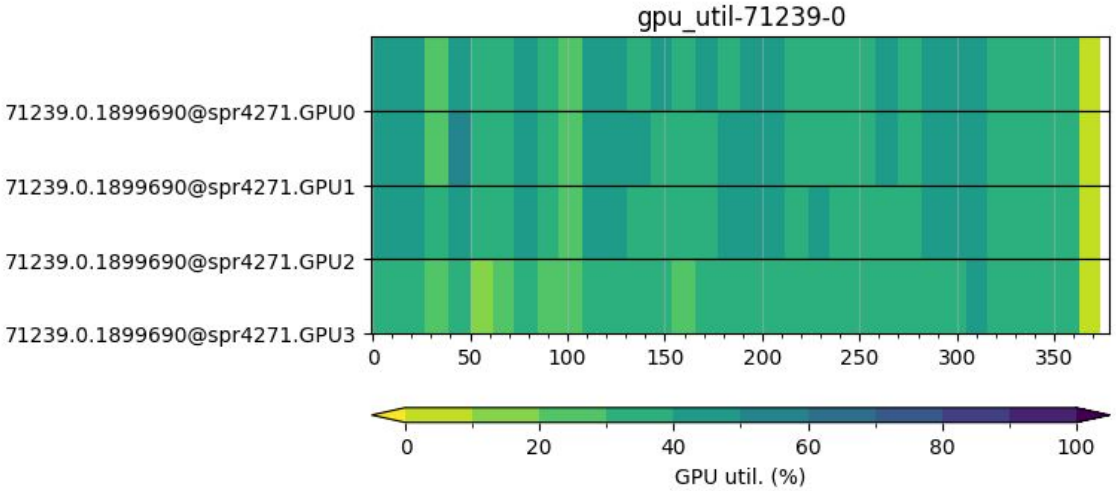
- The goal is to improve the system utilization
- It’s complementary to dynamic energy optimization
- It’s extensible with new powerful rules

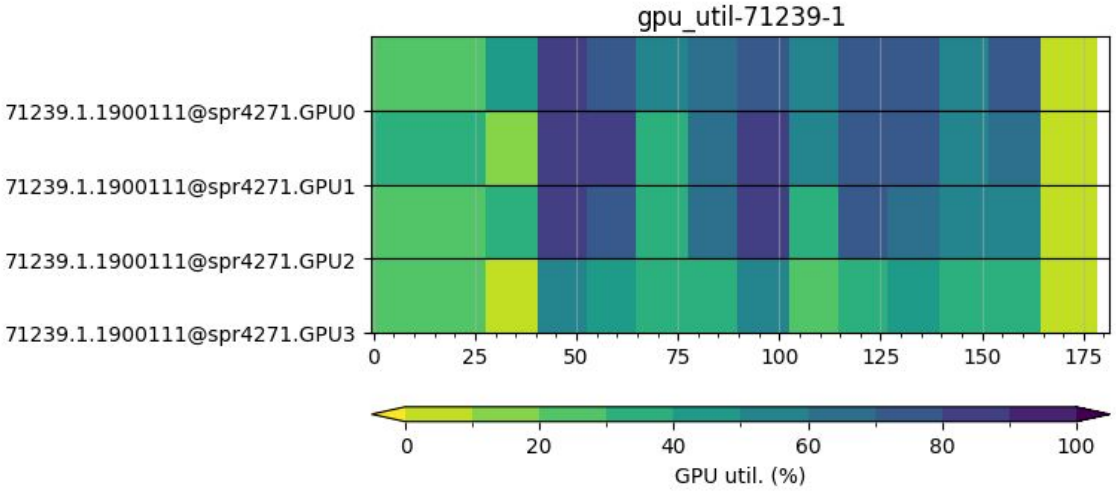
FROM

TO


gpu_power-71239-0
71239.0.1899690@spr4271.GPU0
71239.0.1899690@spr4271.GPU1
71239.0.1899690@spr4271.GPU2
71239.0.1899690@spr4271.GPU3
0 50 100 150 200 250 300 350
100 150 200 250 300 350 400 450
GPU Power (W)


gpu_power-71239-1
71239.1.1900111@spr4271.GPU0
71239.1.1900111@spr4271.GPU1
71239.1.1900111@spr4271.GPU2
71239.1.1900111@spr4271.GPU3
0 25 50 75 100 125 150 175
100 150 200 250 300 350 400 450
GPU Power (W)

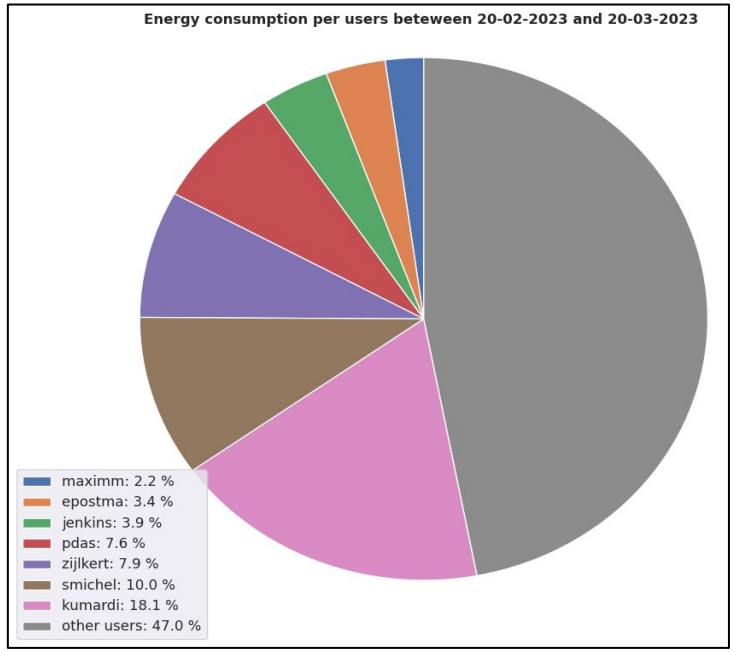

gpu_util-71239-0
71239.0.1899690@spr4271.GPU0
71239.0.1899690@spr4271.GPU1
71239.0.1899690@spr4271.GPU2
71239.0.1899690@spr4271.GPU3
0 50 100 150 200 250 300 350
0 20 40 60 80 100
GPU util. (%)


gpu_util-71239-1
71239.1.1900111@spr4271.GPU0
71239.1.1900111@spr4271.GPU1
71239.1.1900111@spr4271.GPU2
71239.1.1900111@spr4271.GPU3
0 25 50 75 100 125 150 175
0 20 40 60 80 100
GPU util. (%)

EAR system analytics for Monitoring

A data-analysis tool that generates detailed **cluster** and **workload reports** for a given period.

- **Resource consumption** metrics:
  - average power, power over the time, max and min values, etc
  - average temperature, …
- **Workload characteristics**
  - CPU or Memory boundness
  - GFlops reported % vs peaks
  - Energy consumption per users, groups, etc..
- **Anomaly statistics**: (next release)
  - number of anomaly (power, temperature)
  - anomaly distributions (processors, nodes, etc…)



And then …is it all already done?

No :)

- Always new architectures, devices, use cases, schedulers, requirements…
- What we already have
  - Software used in production
  - Homogeneous for users and administrators
  - Homogeneous metrics
  - Cooperative with other tools/products “as much as possible”
- What can we do more?
  - More data center mgt (cooling)
  - More scheduler interaction
  - More utilization of analytics

EAR
Energy Aware Runtime

Thanks for attending!

Barcelona Supercomputing Center
Centro Nacional de Supercomputación